Taking the Difference Between Leisure Time and Workdays Into Account to Improve Virtual Coaching

Talko Dijkhuis[1,2], Johan Blok[1], Hugo Velthuijsen[1], Koen Lemmink [2]
[1]Hanze University of Applied Sciences, Institute for Communication, Media & IT, Groningen, The Netherlands
[2]University of Groningen, Center for Human Movement Sciences, University Medical Center Groningen, Groningen, The Netherlands
[1]{t.b.dijkhuis, j.blok, h.velthuijsen}@pl.hanze.nl
[2]{t.b.dijkhuis, k.a.p.m.lemmink}@umcg.nl

***Abstract*—A sedentary lifestyle is a cause for many health problems. To motivate employees to a healthier lifestyle and to physical activity, the Hanze University initiated a health promotion program, including a fortnightly coaching on lifestyle and activity. An activity tracker was used to monitor the participants' steps. However the used activity tracker doesn't provide timely personalized coaching. In this paper, we investigate the possibility of enhancing timely personalized virtual coaching. Therefore we investigated the manner in which the predictability of physical activity of a participant during the day can be improved. We focussed on the individual differences as well as the difference in activity and the circadian rhythm between free time and workdays. Exploring the data of the experiment the collected step count data was used to examine whether there was a significant difference between leisure time, such as weekends, holidays, and workdays. Taking the findings of this investigation into account, we augmented individual machine learning models. The training of algorithms per participant in combination with time sliced datasets improved the accuracy of the prediction during the day of one meeting his or her daily goal. The use of personalized prediction models, applied machine learning, and the consideration of the difference between leisure time and workdays, will become a valuable and viable addition to a virtual coaching system helping the participants in achieving a healthy lifestyle and enabling a virtual coaching system that intervenes at the appropriate points in time.**

Keywords–monitoring physical activity; circadian rhythm; inference statistics; machine learning; intervention.

I. INTRODUCTION

An unhealthy lifestyle with insufficient daily physical activity shortens life expectancy. Insufficient physical activity is associated with 5.3 million deceases globally in 2008 [1]. Contrarily sufficient physical activity is related to a reduced risk of metabolic syndrome [2], cardiovascular disease [3] and mortality [4]. To promote a healthy lifestyle and physical activity during the workday, the Hanze University of Applied Sciences (HUAS) initiated a health promotion program called -in Dutch- 'Het Gezonde Nieuwe Werken' (HGNW). This initiative contained a focus on the improvement of physical activity. Participants received an activity tracker to increase the awareness of their daily progress in achieving their goals in terms of numbers of steps. The daily feedback of the activity tracker was complemented with a fortnightly coaching session on lifestyle and physical activity. However, the feedback of the activity tracker and its platform didn't provide the participant with timely personalised feedback. Furthermore, current activity trackers do not provide a personalized probability of reaching the daily goal or take the difference between leisure time and workdays into account, therefore lacking an ability to adapt timing of coaching interventions to optimal points in time. Leisure time and workdays are known to show different patterns in activity [5] as well as a different circadian rhythm [6]. To enable a more personalized virtual coach these differences have to be taken into account.

In this paper, we investigate the influence of the circadian pattern of leisure time, workdays and holidays on the physical activity pattern during the day. In addition we investigate whether the level of activity depends on leisure and working time. Pattern differences due to these factors provide an opportunity to implement personalized automated intervention strategies for an individualized virtual coach.

The remainder of this paper is organized as follows. The first section introduces the state of the art on measuring activity levels, and the use of machine learning for monitoring. Subsequently, we describe the study on health promotion at HUAS, the collected dataset on daily physical activity of the participants, the method of statistical analysis of the difference between week, workweek, weekend, holiday and bank holiday, and measuring the performance of the trained machine learning models. In the third section we present the results of the statistical analysis and the performance of trained machine learning models. The conclusion on the results and a short discussion on future work completes the paper.

II. STATE OF THE ART

Activity trackers provide a measure for the number of steps and enable monitoring progress during the day and over time. Adding an activity tracker on steps to physical therapy or counselling was effective in some groups [7] [8]. The collection of step data is not only effective for therapy or counselling, it is also an intervention mechanism in itself [9]. Only the fact of using an activity tracker could motivate physical activity and improvement of health [10]. To improve on physical activity in combination with activity tracking monitoring, coaching is helpful. Effectiveness of (e)Coaching depends on timeliness and on personal contextual information in combination with actionable insights [11]. In other words the participant needs to receive the information and the advice when it is most relevant. To the best of our knowledge no studies exist on the use of activity trackers in combination with the circadian pattern and machine learning algorithms to establish *individualized* models or studies on *individualized* models used in virtual coach systems on monitoring activity helping the participant to improve his or her physical behaviour.

 ISBN: 978-1-61208-688-0

III. METHODS

In this section, we present the study design of the HNGW, the data set we analysed, the statistical method to identify the differences in physical activity between leisure time and work week, and the training of algorithms with time sliced data and the methods used for analysis of the accuracy of the trained models.

A. Study design

The study data stems from the HNGW project. Forty eight healthy employees were recruited from the HUAS. The 48 participants were divided according to age, gender, BMI, and baseline self-reported health prior to being randomized into two groups. Group A followed a twelve-week health promotion intervention; the other group, group B, served first as a control group and thereafter also received the twelve-week health promotion intervention. Only the intervention period was used to study. The outcome measures included, among other values, the daily steps. The daily steps were measured with the Fitbit Flex, which is known to be a trustworthy and valid activity tracker for step count [12] and suitable for health promotion programs [13].

B. Data set

We prepared the individual participants available minute step data to investigate the difference between the activity patterns on the weekends, the holidays, the bank holidays and the workdays, and the corresponding activity patterns. We followed a step-by-step approach. First, we performed a data pre-processing step to remove the incomplete records from the data set. We also eliminated all records per day whenever no step was gathered during that day. Second, we augmented an hourly summarised data set per participant with new derived variables representing:

1) the year (2014-2015)
2) the week of the year (range 0-52)
3) the day of the week (range 0 - 6)
4) the hour of day (range 0 - 23)
5) the steps per hour
6) the cumulative sum of the steps per hour

Third, the work week is defined as the weekdays Monday till Friday. Fourth, a weekend is defined as Saturday and Sunday. Fifth, because all employees work on at an university of applied sciences, they have the same holidays and information on the holidays was added to the data set. Sixth, the bank holidays were identified and added. For the column the average number of steps per participant per day the amount of steps between 6:00AM and 11:00PM was considered. The average number of steps per participant per day the amount of steps column was regarded as a threshold per day in order to determine the binary outcome column. Finally, we constructed individual binary outcome variables based on the threshold of the time slices work week, weekend, holiday, bank holiday, and week.

C. Statistical Analysis

To determine whether there were differences in level of activity and moment of activity, we fitted a regression model and used the ANOVA analysis in conjunction with the Tukey HSD post-hoc analysis to identify possible correlations and differences between the hourly datasets whole week, work week, weekend, holiday and bank holiday. We reject the null hypothesis when the mean of the steps per hours is equal between the different time slices.

The Random Forest algorithm was chosen as the algorithm to enable prediction on whether a participant would meet his daily threshold. Random Forest is known as one of the most accurate algorithms predicting the participant meeting his or her goal [14]. The Random Forest algorithm was trained for the individual participant and hourly time sliced datasets. The techniques cross validation and parameter tuning were used to optimize the trained model. To investigate the performance of individualized hourly work week, weekend, holiday and bank holiday models, we compared these predictions with the predictions of a baseline model. The baseline model was trained per participant on the whole data set without distinction between the time slices work week, weekend, holiday or bank holiday and then the baseline model predicted adopting the four time sliced datasets leading to four sets of predictions. Next the four models for the work week, weekend, holiday and bank holiday predicted adopting the four time sliced datasets. The confusion matrix method was used to classify the difference between the predicted value and the actual value of both the baseline model and the time sliced models. A confusion matrix provides an overview of the true positives (TP; a predicted a 'true' and the actual data contained a 'true'), true negatives (TN; the model predicted a 'false' and the actual data was a 'false'), false positives (FP; the model predicted a 'true' label, but the actual data was a 'false'), and false negatives (FN; the model predicted a 'false' label, but the data was 'true') of a model. The confusion matrix served as a basis for the calculation of the performance measure F1-score and the accuracy [15].

The F1-score and accuracy are calculated for each model on group level, both metrics have a range of zero to one, where one is the best score. To calculate the F1-score, two other metrics known as the precision and the recall are used. Precision is the proportion of the true positives and the false negatives, and is calculated as:

$$\frac{TP}{(TP+FN)}$$

Recall is the true positive rate, which is calculated as follows:

$$\frac{TP}{(TP+FP)}$$

Using precision and recall, the F1-score is calculated as:

$$2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

To calculate the accuracy is a metric to determine the nearness of the prediction to the true value. A value of the accuracy close to one indicates the best performance. It calculates the ratio between the correctly classified cases and all cases as:

$$\frac{TP+TN}{TP+TN+FP+FN}$$

The personalized work week, weekend, holiday and bank holiday models were studied on the performance in comparison with the baseline week model.

IV. RESULTS

The ANOVA test and the Tukey HSD post-hoc analysis identified significant differences between the hourly level of activity and the time sliced datasets. The ANOVA test found a significant difference in activity level ($p<0.001$) between the work week, the weekend, the holiday and the bank holiday at six, seven, eight o'clock. The participants were less active in the beginning of the day. Fig. 1 is an illustration of the difference of level of activity per hour of the day between the work week and the weekend.

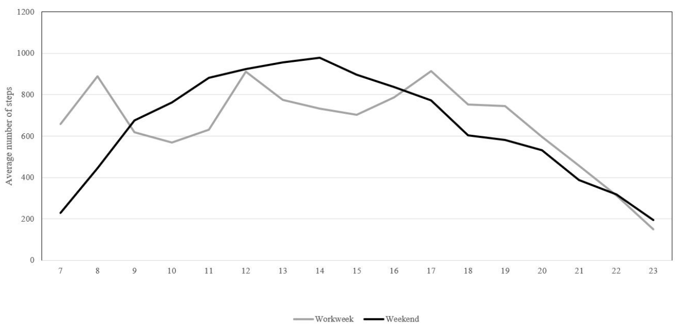

Figure 1. The difference in pattern of number of steps per hour during the work week versus the weekend.

The Tukey HSD post-hoc analysis showed where there are significant differences between the time sliced datasets. The majority of the datasets (21 out of 30) showed significant differences at six, seven, and eight o' clock. In the remainder of the day the differences between the different time sliced datasets were not significant.

Table I is an example of the Tukey HSD post-hoc analysis representing the seven o'clock results.

TABLE I. MULTIPLE COMPARISON OF MEANS - TUKEY HSD, ALFA=0.05 EXAMPLE

Dataset 1	*Dataset 2*	*Hour*	*Mean difference*	*Reject*
all week	bank holiday	7	-225	True
all week	holiday	7	-167	True
all week	weekend	7	-316	True
all week	work week	7	165	True
bank holiday	holiday	7	88	False
bank holiday	weekend	7	-60	False
bank holiday	work week	7	421	True
holiday	weekend	7	-148	True
holiday	work week	7	333	True
weekend	work week	7	481	True

The personalized Random Forest models based on the hourly datasets proved to show a better F1-score and accuracy then the baseline week model. Table II an Table III represent the improvement of the performance when using the work week, weekend, holiday and bank holiday model instead of the baseline weekmodel.

TABLE II. COMBINATIONS OF TIME SLICE BASED MODELS AND THEIR F1-SCORE.

Model 1	*F1-score*	*Model 2*	*F1-score*	*Difference*
week model	0.59	weekend model	0.86	48%
week model	0.58	holiday model	0.71	22%
week model	0.57	bank holiday model	0.68	19%
week model	0.81	work week model	0.96	18%

V. CONCLUSION

The level of activity in the morning is influenced by the difference between leisure time and work time. The individualized time sliced models improved the F1-score and accuracy in comparison to the non-time sliced week model. It is recommended to construct time sliced models per individual to improve the performance of the individualized models. In the future, contextual data that influences physical activity, like weekly non-workdays, sports, and celebrations, may be taken into account. Another possible future direction is to individualize interventions that allow for more personalized coaching. The individualization of the predictive models enables automated personalized, contextualized, and timely coaching. The results of this paper will be applied in the preventive eHealth virtual coach platform as suggested by Blok et al. [16].

TABLE III. COMBINATIONS OF TIME SLICE BASED MODELS AND THEIR ACCURACY.

Model 1	*Accuracy*	*Model 2*	*Accuracy*	*Difference*
week model	0.89	weekend model	0.92	2.9%
week model	0.89	holiday model	0.89	0%
week model	0.76	bank holiday model	0.98	28%
week model	0.81	work week model	0.96	18%

ACKNOWLEDGMENT

The authors thank the Hanze University Health Program, especially M. van Ittersum, for providing the physical activity data of the Health Program, all the participants in the experiment, and M. Aiello for suggestions on the structure of the paper.

REFERENCES

[1] I. Min-Lee et al., "Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy," The Lancet, vol. 380, no. 9838, 2012, pp. 219–229. [Online]. Available: http://dx.doi.org/10.1016/S0140-6736(12)61031-9

[2] D. Zhang et al., "Leisure-time physical activity and incident metabolic syndrome: a systematic review and dose-response meta-analysis of cohort studies," pp. 36–44, oct 2017. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/28927737https://linkinghub.elsevier.com/retrieve/pii/S002604951730207X

[3] S. A. Lear et al., "The effect of physical activity on mortality and cardiovascular disease in 130 000 people from 17 high-income, middle-income, and low-income countries: the PURE study," www.thelancet.com, vol. 390, 2017. [Online]. Available: http://dx.doi.org/10.1016/

[4] D. D. Carter, K. Robinson, J. Forbes, and S. Hayes, "Experiences of mobile health in promoting physical activity: A qualitative systematic review and meta-ethnography," PLOS ONE, vol. 13, no. 12, dec 2018, p. e0208759. [Online]. Available: http://dx.plos.org/10.1371/journal.pone.0208759

[5] C. E. Matthews et al., "Sources of Variance in Daily Physical Activity Levels in the Seasonal Variation of Blood Cholesterol Study," American Journal of Epidemiology, vol. 153, no. 10, 2001, pp. 987–995.

[6] S.-J. Paine and P. H. Gander, "Explaining ethnic inequities in sleep duration: a cross-sectional survey of Mori and non-Mori adults in New Zealand." Sleep health, vol. 2, no. 2, jun 2016, pp. 109–115. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/28923252

[7] H. J. de Vries, T. J. Kooiman, M. W. van Ittersum, M. van Brussel, and M. de Groot, "Do activity monitors increase physical activity in adults with overweight or obesity? A systematic review and meta-analysis," Obesity, vol. 24, no. 10, oct 2016, pp. 2078–2091. [Online]. Available: http://doi.wiley.com/10.1002/oby.21619

[8] M. Miyauchi, M. Toyoda, N. Kaneyama, H. Miyatake, E. Tanaka, M. Kimura, T. Umezono, and M. Fukagawa, "Exercise Therapy for Management of Type 2 Diabetes Mellitus: Superior Efficacy of Activity Monitors over Pedometers," Journal of Diabetes Research, vol. 2016, sep 2016, pp. 1–7. [Online]. Available: https://www.hindawi.com/journals/jdr/2016/5043964/

[9] L. A. Cadmus-Bertram, B. H. Marcus, R. E. Patterson, B. A. Parker, and B. L. Morey, "Randomized Trial of a Fitbit-Based Physical Activity Intervention for Women." American journal of preventive medicine, vol. 49, no. 3, sep 2015, pp. 414–8. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/26071863http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4993151

[10] Z. H. Lewis, E. J. Lyons, J. M. Jarvis, and J. Baillargeon, "Using an electronic activity monitor system as an intervention modality: A systematic review," BMC Public Health, vol. 15, no. 1, dec 2015, p. 585. [Online]. Available: http://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-015-1947-3

[11] M. Gerdes, S. Martinez, and D. Tjondronegoro, "Conceptualization of a Personalized eCoach for Wellness Promotion," in Proceedings of ACM 11th EAI International Conference on Pervasive Computing. Barcelona: Association for Computing Machinery, 2017.

[12] T. J. M. Kooiman, M. L. Dontje, S. R. Sprenger, W. P. Krijnen, C. P. van der Schans, and M. de Groot, "Reliability and validity of ten consumer activity trackers," BMC Sports Science, Medicine and Rehabilitation, vol. 7, no. 1, dec 2015, p. 24. [Online]. Available: http://bmcsportsscimedrehabil.biomedcentral.com/articles/10.1186/s13102-015-0018-5

[13] J. Wang, R. Chen, X. Sun, M. F. H. She, and Y. Wu, "Recognizing human daily activities from accelerometer signal," Procedia Engineering, vol. 15, 2011, pp. 1780–1786. [Online]. Available: http://dx.doi.org/10.1016/j.proeng.2011.08.331

[14] T. B. Dijkhuis, F. J. Blaauw, M. W. van Ittersum, H. Velthuijsen, and M. Aiello, "Personalized Physical Activity Coaching: A Machine Learning Approach," Sensors, vol. 18, no. 2, feb 2018, p. 623. [Online]. Available: http://www.mdpi.com/1424-8220/18/2/623

[15] S. Schoeppe et al., "Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review," International Journal of Behavioral Nutrition and Physical Activity, vol. 13, no. 1, dec 2016, p. 127. [Online]. Available: http://ijbnpa.biomedcentral.com/articles/10.1186/s12966-016-0454-y

[16] J. Blok, A. Dol, and T. Dijkhuis, "Toward a Generic Personalized Virtual Coach for Self-management : a Proposal for an Architecture," no. c, 2017, pp. 105–108.